Mutations driving CLL and their evolution in progression and relapse

Landau, Tausch, Taylor-Weiner et al.

TABLE OF CONTENTS

Supplementary Table 1: Patients characteristics of the 278 patients that provided samples as part of the GCLLSG-CLL8 trial.

Characteristic	Sample not available, N (%)	Sample available, N (%)	*P* value	Total N (%)
Patient number	**539**	**278**		**817**
Treatment arm				
FC	266 (49.4)	143 (51.4)	0.605	409
FCR	273 (50.6)	135 (48.6)		408
Median age (range) (years)	61 (30-78)	60 (31-81)	0.115	61 (30-81)
Age group (years)				
≥ 65	174 (32.3)	71 (25.5)	0.053	245 (30.0)
≥ 70	60 (11.6)	21 (7.6)	0.11	81 (9.9)
Sex				
Female	140 (26.0)	70 (25.2)	0.866	210 (25.7)
Male	399 (74.0)	208 (74.8)		607 (74.3)
Binet stage	**539**	**278**		**817**
A	25 (4.6)	15 (5.4)	0.442	40 (4.9)
B	352 (65.3)	170 (61.2)		522 (63.9)
C	159 (29.5)	93 (33.5)		252 (30.8)
Missing	3 (0.6)	0 (0.0)		3 (0.4)
ECOG performance status	**524**	**261**		**785**
0	311 (59.4)	136 (52.1)	0.056	447 (56.9)
> 0	213 (40.6)	125 (47.9)		338 (43.1)
Presence of B symptoms	**536**	**277**		**813**
No	295 (55.0)	154 (55.6)	0.882	449 (55.2)
Yes	241 (45.0)	123 (44.4)		364 (44.8)
***IGHV* mutational status**	**358**	**264**		**622**
Unmutated	234 (65.4)	158 (59.8)	0.179	392 (63.0)
Mutated	124 (34.6)	106 (40.2)		230 (37.0)
Deletion 17p by FISH	**358**	**263**		**621**
No	320 (89.4)	250 (95.1)	0.012	570 (91.8)
Yes	38 (10.6)	13 (4.9)		51 (8.2)

Deletion 11q by FISH		**263**		**621**
No	274 (76.5)	194 (73.8)	0.452	468 (75.4)
Yes	84 (23.5)	69 (26.2)		153 (24.6)
Trisomy 12 by FISH	**356**	**262**		**618**
No	318 (89.3)	226 (86.3)	0.261	544 (88.0)
Yes	38 (10.7)	36 (13.7)		74 (12.0)
Deletion 13q by FISH	**356**	**261**		**617**
No	151 (42.4)	116 (44.4)	0.623	267 (43.3)
Yes	205 (57.6)	145 (55.6)		350 (56.7)
Genetic classification according to hierarchical model by FISH	**355**	**261**		**616**
Del(17p)	38 (10.7)	13 (5.0)	0.056	51 (8.3)
Del(11q)	75 (21.1)	67 (25.7)		142 (23.1)
Trisomy 12	30 (8.5)	31 (11.9)		61 (9.9)
No abnormalities	81 (22.8)	57 (21.8)		138 (22.4)
Del(13q)	131 (36.9)	93 (35.6)		224 (36.4)
s- β_2m (mg/l)	**343**	**248**		**591**
Median (range)	2.9 (0.7-10.2)	2.8 (0.9-8.0)	0.076	2.9 (0.7-8.0)
< 3.5	223 (65.0)	171 (69.0)	0.332	394 (66.7)
≥ 3.5	120 (35.0)	77 (31.0)		197 (33.3)
ZAP-70 expression	**146**	**143**		**289**
≤ 20	89 (61.0)	86 (60.1)	0.887	175 (60.6)
> 20	57 (39.0)	57 (39.9)		114 (39.4)
CD38 expression	**458**	**258**		**716**
≤ 30	308 (67.2)	170 (65.9)	0.711	478 (66.8)
> 30	150 (32.8)	88 (34.1)		238 (33.2)
Response to treatment	**539**	**278**		
Response	441 (81.8)	256 (92.1)	<0.01	697 (85.3)
Non-response	52 (9.6)	10 (3.6)		120 (14.7)
Missing	46 (8.5)	12 (4.3)		
CR	156 (28.9)	96 (34.5)	0.110	252 (30.8)
Non-CR/missing	383 (71.1)	182 (65.5)		565 (69.2)

* Responder versus non-response/missing

Supplementary Table 5: Comparison between somatic copy number variations (sCNVs) detection with WES vs. FISH cytogenetics

	Deletion 17p by FISH	
	No	Yes
Deletion 17p by WES detection, N (%)	**250**	**13**
No	249 (99.6)	0 (0.0)
Yes	1 (0.4)	13 (100.0)

	Deletion 11q by FISH	
	No	Yes
Deletion 11q by WES detection, N (%)	**194**	**69**
No	189 (97.4)	3 (4.3)
Yes	7 (3.6)	66 (95.7)

	Trisomy 12 by FISH	
	No	Yes
Trisomy 12 by WES detection, N (%)	**226**	**36**
No	225 (99.6)	1 (2.8)
Yes	1 (0.4)	35 (97.2)

	Deletion 13q by FISH	
	No	Yes
Deletion 13q by WES detection, N (%)	**116**	**145**
No	116 (100.0)	21 (14.5)
Yes	0 (0.0)	124 (85.5)

Supplementary Table 7: Temporal order of somatic mutation acquisitions – classifying drivers as early vs. late.

These table includes all driver events (recurrent sCNVs and candidate CLL gene non-silent mutations), classifying the driver events based on the relative enrichment of out-degrees vs. in-degrees as early (Q<0.2 and number of out-degrees > in-degrees), late (Q<0.2 and number of out-degrees < in-degrees) and intermediate or not powered (Q>0.2, Inter./not powered). Out-degrees are defined as instances in which the driver event is clonal and found in the same CLL with another driver event that is subclonal. In-degrees are defined as instances in which the driver event is subclonal and found in the same CLL with another driver event that is clonal.

Table 7a: Results in n = 501 treatment naive patients (4 patients with unknown status of prior therapy were excluded from the analysis as well):

Driver event	*Q*-value	occur-rences	in-degrees	out-degrees	classification
del13q	1.22E-23	233	35	183	Early
tri12	9.44E-22	67	1	81	Early
ATM	2.94E-13	76	89	14	Late
BIRC3	2.05E-07	15	27	0	Late
del11q	3.93E-04	103	42	90	Early
del20p	2.24E-03	6	0	13	Early
FBXW7	1.53E-02	10	10	0	Late
MAP2K1	2.39E-02	8	9	0	Late
NRAS	2.39E-02	8	9	0	Late
KRAS	3.49E-02	14	11	1	Late
BAZ2A	3.58E-02	9	8	0	Late
MYD88	3.58E-02	14	0	8	Early
CARD11	4.64E-02	7	10	1	Late
MGA	4.64E-02	15	16	4	Late
ZMYM3	4.74E-02	10	12	2	Late
TP53	6.90E-02	29	26	11	Late
amp2p	9.54E-02	45	44	25	Late
del6q21	1.60E-01	16	16	6	Late
DYRK1A	1.93E-01	7	7	1	Late
FAM50A	1.93E-01	5	7	1	Late
CHEK2	3.12E-01	5	1	6	Inter./not powered
TRAF3	3.12E-01	4	4	0	Inter./not powered
BRAF	3.35E-01	19	12	5	Inter./not powered
IRF4	3.35E-01	10	9	3	Inter./not powered
NOTCH1	3.38E-01	37	25	15	Inter./not powered
amp8q	3.53E-01	12	10	4	Inter./not powered

MED12	3.53E-01	7	7	2	Inter./not powered
TRAF2	3.53E-01	6	7	2	Inter./not powered
PTPN11	4.15E-01	6	5	1	Inter./not powered
PIM1	4.44E-01	2	3	0	Inter./not powered
tri19	4.44E-01	6	0	3	Inter./not powered
ASXL1	4.82E-01	5	6	2	Inter./not powered
ELF4	4.82E-01	7	6	2	Inter./not powered
RPS15	5.29E-01	21	16	10	Inter./not powered
SF3B1	5.32E-01	103	72	60	Inter./not powered
NXF1	5.67E-01	7	4	1	Inter./not powered
POT1	5.67E-01	33	20	27	Inter./not powered
GNB1	6.56E-01	5	5	2	Inter./not powered
HIST1H1E	7.16E-01	7	3	6	Inter./not powered
BRCC3	7.55E-01	5	7	4	Inter./not powered
IKZF3	7.79E-01	11	5	8	Inter./not powered
del18p	8.24E-01	12	10	7	Inter./not powered
EGR2	8.67E-01	16	10	13	Inter./not powered
del17p	9.10E-01	24	15	18	Inter./not powered
FUBP1	9.61E-01	9	8	6	Inter./not powered
CHD2	9.61E-01	23	9	7	Inter./not powered
XPO1	9.95E-01	22	13	15	Inter./not powered
BCOR	1.00E+00	10	7	6	Inter./not powered
IGLL5	1.00E+00	11	6	7	Inter./not powered
DDX3X	1.00E+00	10	4	4	Inter./not powered
del8p	1.00E+00	16	10	10	Inter./not powered
EWSR1	1.00E+00	4	1	2	Inter./not powered
HIST1H1B	1.00E+00	4	2	2	Inter./not powered
SAMHD1	1.00E+00	10	6	6	Inter./not powered
XPO4	1.00E+00	7	5	5	Inter./not powered

Table 7b: Results in n = 229 treatment naive patients with *IGHV* unmutated CLL:

Driver event	*Q*-value	occurr-ences	in-degrees	out-degrees	classification
tri12	3.61E-16	36	0	58	Early
del13q	7.78E-15	85	20	110	Early
ATM	3.90E-07	56	58	12	Late
BIRC3	1.98E-04	10	17	0	Late
del20p	2.54E-03	5	0	13	Early
del11q	2.71E-03	78	28	63	Early
MGA	3.85E-03	13	15	1	Late
FBXW7	2.54E-02	7	9	0	Late
NRAS	9.03E-02	6	7	0	Late
ASXL1	1.02E-01	6	9	1	Late
KRAS	1.02E-01	12	9	1	Late
amp2p	1.04E-01	33	34	17	Late
NOTCH1	1.08E-01	20	19	7	Late
BAZ2A	1.08E-01	7	6	0	Late
MAP2K1	1.08E-01	5	6	0	Late
IRF4	1.19E-01	10	8	1	Late
MED12	1.19E-01	7	8	1	Late
TP53	1.51E-01	17	19	8	Late
ELF4	1.70E-01	5	5	0	Early
ZMYM3	1.70E-01	8	9	2	Late
CHEK2	1.74E-01	4	1	7	Early
BRAF	1.82E-01	16	12	4	Late
del8p	2.60E-01	8	4	11	Inter./not powered
GNB1	2.60E-01	3	4	0	Inter./not powered
PTPN11	2.60E-01	4	4	0	Inter./not powered
CARD11	4.06E-01	4	5	1	Inter./not powered
CHD2	4.06E-01	4	5	1	Inter./not powered
del18p	4.06E-01	6	1	5	Inter./not powered
del17p	4.12E-01	16	9	16	Inter./not powered
IGLL5	4.33E-01	2	3	0	Inter./not powered
DYRK1A	4.85E-01	6	6	2	Inter./not powered
del6q21	5.40E-01	13	11	6	Inter./not powered
NXF1	5.91E-01	7	4	1	Inter./not powered
XPO4	6.67E-01	5	2	5	Inter./not powered
POT1	6.67E-01	20	15	20	Inter./not powered
HIST1H1B	6.67E-01	1	2	0	Inter./not powered
HIST1H1E	6.67E-01	1	2	0	Inter./not powered
PIM1	6.67E-01	1	2	0	Inter./not powered

TRAF3	6.67E-01	3	2	0	Inter./not powered
EGR2	8.54E-01	14	9	12	Inter./not powered
XPO1	8.54E-01	20	10	13	Inter./not powered
RPS15	8.54E-01	19	14	11	Inter./not powered
BRCC3	8.79E-01	4	5	3	Inter./not powered
IKZF3	8.91E-01	8	4	6	Inter./not powered
DDX3X	1.00E+00	8	3	4	Inter./not powered
SAMHD1	1.00E+00	7	3	4	Inter./not powered
SF3B1	1.00E+00	60	41	42	Inter./not powered
BCOR	1.00E+00	8	7	6	Inter./not powered
amp8q	1.00E+00	5	1	2	Inter./not powered
FAM50A	1.00E+00	1	0	1	Inter./not powered
FUBP1	1.00E+00	6	5	5	Inter./not powered
TRAF2	1.00E+00	4	3	2	Inter./not powered

Table 7c: Results in n = 245 treatment naive patients with *IGHV* mutated CLL:

Driver event	*Q*-value	occurre-nces	in-degrees	out-degrees	classification
del13q	1.56E-06	130	10	53	Early
ATM	7.31E-06	19	22	0	Late
tri12	7.31E-06	23	0	22	Early
BIRC3	2.81E-03	7	13	0	Late
MYD88	1.44E-01	13	0	7	Early
NOTCH1	2.40E-01	5	6	0	Inter./not powered
tri19	4.11E-01	5	0	5	Inter./not powered
amp8q	5.23E-01	5	6	1	Inter./not powered
CARD11	5.23E-01	3	4	0	Inter./not powered
FAM50A	5.23E-01	3	4	0	Inter./not powered
TRAF2	5.23E-01	2	4	0	Inter./not powered
FUBP1	8.85E-01	3	3	0	Inter./not powered
XPO1	8.85E-01	3	3	0	Inter./not powered
TP53	1.00E+00	12	5	2	Inter./not powered
BAZ2A	1.00E+00	2	2	0	Inter./not powered
BRCC3	1.00E+00	1	2	0	Inter./not powered
del18p	1.00E+00	1	0	2	Inter./not powered
del6q21	1.00E+00	1	2	0	Inter./not powered
HIST1H1B	1.00E+00	3	0	2	Inter./not powered
KRAS	1.00E+00	2	2	0	Inter./not powered
XPO4	1.00E+00	1	2	0	Inter./not powered
IGLL5	1.00E+00	9	3	6	Inter./not powered
del8p	1.00E+00	7	3	1	Inter./not powered
HIST1H1E	1.00E+00	3	1	3	Inter./not powered
SF3B1	1.00E+00	38	23	19	Inter./not powered
del11q	1.00E+00	20	13	10	Inter./not powered
CHD2	1.00E+00	19	4	6	Inter./not powered
amp2p	1.00E+00	2	1	2	Inter./not powered
DDX3X	1.00E+00	2	1	0	Inter./not powered
del17p	1.00E+00	8	3	2	Inter./not powered
del20p	1.00E+00	1	0	1	Inter./not powered
DYRK1A	1.00E+00	1	1	0	Inter./not powered
ELF4	1.00E+00	2	1	2	Inter./not powered
EWSR1	1.00E+00	3	1	2	Inter./not powered
FBXW7	1.00E+00	2	1	0	Inter./not powered
GNB1	1.00E+00	2	1	2	Inter./not powered
IKZF3	1.00E+00	1	0	1	Inter./not powered
MGA	1.00E+00	3	2	3	Inter./not powered

NRAS	1.00E+00	2	1	0	Inter./not powered
PIM1	1.00E+00	1	1	0	Inter./not powered
POT1	1.00E+00	11	4	5	Inter./not powered
PTPN11	1.00E+00	2	1	1	Inter./not powered
RPS15	1.00E+00	1	1	0	Inter./not powered
SAMHD1	1.00E+00	3	3	2	Inter./not powered
TRAF3	1.00E+00	1	1	0	Inter./not powered
ZMYM3	1.00E+00	1	1	0	Inter./not powered

Supplementary Table 8: Temporal order of somatic mutation acquisitions – pairwise data.

These tables includes all pairs of driver events (d1, d2) that had at least 5 cases in which the two drivers were detected in the same CLL sample, but one of the drivers is clonal and the other is subclonal. A two-tailed binomial test is performed to test whether the pairings are found to be in one order more frequently than the other (i.e., d1 → d2 > d2 → d1). A multi-hypothesis correction is then applied and the table lists all hypotheses tested.

Table 8a: Pairing in n = 501 treatment naive patients (4 patients with unknown status of prior therapy were excluded from the analysis as well):

Ordering	*P*-value	*Q*-value	No. orderings (clonal → subclonal)	No. opposite orderings (subclonal → clonal)
del13q->ATM	1.54E-08	3.53E-07	31	1
del11q->ATM	2.10E-05	2.41E-04	20	1
del13q->amp2p	6.10E-05	4.68E-04	15	0
del13q->SF3B1	3.24E-04	1.86E-03	27	6
del11q->amp2p	9.77E-04	4.49E-03	14	1
tri12->BIRC3	3.91E-03	1.50E-02	9	0
tri12->del11q	3.12E-02	1.03E-01	6	0
del11q->BIRC3	6.25E-02	1.20E-01	5	0
del13q->del6q21	6.25E-02	1.20E-01	5	0
del13q->MGA	6.25E-02	1.20E-01	5	0
tri12->FBXW7	6.25E-02	1.20E-01	5	0
tri12->KRAS	6.25E-02	1.20E-01	5	0
amp2p->ATM	7.03E-02	1.24E-01	7	1
del13q->del11q	9.31E-02	1.53E-01	16	7
del13q->POT1	1.25E-01	1.80E-01	6	1
del13q->TP53	1.25E-01	1.80E-01	6	1
del11q->SF3B1	2.67E-01	3.61E-01	9	4
del13q->del17p	3.75E-01	4.54E-01	4	1
del13q->NOTCH1	3.75E-01	4.54E-01	4	1
SF3B1->ATM	5.08E-01	5.84E-01	6	3
del17p->TP53	6.88E-01	7.53E-01	4	2
XPO1->del13q	1.00E+00	1.00E+00	4	3
POT1->del11q	1.00E+00	1.00E+00	3	2

Table 8b: Pairing in n = 229 treatment naive patients with *IGHV* unmutated CLL. Patients with unknown *IGHV* status were excluded from the analysis:

Ordering	*P*-value	*Q*-value	No. orderings (clonal → subclonal)	No. opposite orderings (subclonal → clonal)
del13q->ATM	1.45E-04	1.71E-03	17	1
del11q->ATM	2.44E-04	1.71E-03	13	0
del13q->amp2p	9.77E-04	4.56E-03	11	0
del11q->amp2p	6.35E-03	2.11E-02	11	1
del13q->SF3B1	7.54E-03	2.11E-02	15	3
amp2p->ATM	6.25E-02	1.09E-01	5	0
del11q->BIRC3	6.25E-02	1.09E-01	5	0
del13q->MGA	6.25E-02	1.09E-01	5	0
del13q->del11q	3.02E-01	4.69E-01	10	5
del13q->TP53	3.75E-01	5.25E-01	4	1
SF3B1->amp2p	4.53E-01	5.77E-01	5	2
POT1->SF3B1	6.88E-01	8.02E-01	4	2
del11q->SF3B1	1.00E+00	1.00E+00	4	3
SF3B1->ATM	1.00E+00	1.00E+00	4	3

Table 8c: Pairings in n = 245 treatment naive patients with *IGHV* mutated CLL. Patients with unknown *IGHV* status were excluded from the analysis:

Ordering	*P*-value	*Q*-value	No. orderings (clonal → subclonal)	No. opposite ordering (subclonal → clonal)
del13q->ATM	9.77E-04	4.88E-03	11	0
tri12->BIRC3	1.56E-02	3.91E-02	7	0
del13q->SF3B1	9.23E-02	1.54E-01	10	3
del13q->del11q	1.25E-01	1.56E-01	6	1
SF3B1->IGLL5	1.00E+00	1.00E+00	3	2

Supplementary Table 9: Description of patients for which matched pre-treatment and relapse samples were analyzed.

Characteristic	N (%)
All patients with longitudinal data available, N	**59**
Treatment arm	
FC	28 (47.5)
FCR	31 (52.5)
Median age (range) (years)	58 (36-71)
Age group (years)	
≥ 65	9 (15.3)
≥ 70	3 (5.1)
Sex	
Female	13 (22.0)
Male	46 (78.0)
Binet stage	
A	3 (5.1)
B	43 (72.9)
C	13 (22.0)
ECOG performance status	**58**
0	33 (56.9)
> 0	25 (43.1)
***IGHV* mutational status**	**57**
Unmutated	42 (73.7)
Mutated	15 (26.3)
Deletion 17p by FISH	**57**
No	53 (93.0)
Yes	4 (7.0)
Deletion 11q by FISH	**57**
No	35 (61.4)
Yes	22 (38.6)
Trisomy 12 by FISH	**57**

No	53 (93.0)
Yes	4 (7.0)
Deletion 13q by FISH	**57**
No	24 (42.1)
Yes	33 (57.9)
Genetic classification according to hierarchical model by FISH	**57**
Del(17p)	4 (7.0)
Del(11q)	22 (38.6)
Trisomy 12	4 (7.0)
No abnormalities	10 (17.5)
Del(13q)	17 (29.8)
ZAP-70 expression	**31**
≤ 20	17 (54.8)
> 20	14 (45.2)
CD38 expression	**55**
≤ 30	37 (67.3)
> 30	18 (32.7)
Response to treatment	**59**
Response	57 (96.6)
Non-response	2 (3.4)
CR	14 (23.7)
Non-CR	45 (76.3)

METHODS

Human samples: Heparinized blood was obtained from patients enrolled on the prospective, randomized, open-label CLL8 trial[1] before the first cycle of treatment. Sample selection was based on availability, and the baseline characteristics of the cohort of patients studied with WES are largely reflective of the baseline characteristics of the entire CLL8 cohort (***Supplementary Table 1***). All patients had a diagnosed CLL according to WHO criteria confirmed by flow cytometry and were in Binet Stage C or Binet A or B with need for treatment as defined by the study inclusion criteria[1]. Peripheral blood mononuclear cells (PBMC) from patients were isolated by Ficoll density gradient centrifugation. Immuno-magnetic tumor cell enrichment via CD19 was performed on all baseline pretreatment samples (Midi MACS, Miltenyi Biotec, Bergisch Gladbach, Germany) achieving a separation of PBMCs into a CD19-positive tumor sample and a CD19-negative normal sample with a purity of >95% by flow cytometry. For some samples, the source of matched normal tissue was PBMC collected following chemoimmunotherapy when the samples were evaluated by flow cytometry as minimal residual disease-negative (Kiel Laboratory, Germany). Samples were used fresh or cryopreserved (with FBS with10% DMSO) and stored in vapor-phase liquid nitrogen until the time of analysis.

Established CLL prognostic factor analysis: Analyses of genomic aberrations and immunoglobulin heavy-chain variable (*IGHV*) homology were performed in the central reference laboratory of the German CLL Study Group (GCLLSG) in Ulm, Germany. Unmutated *IGHV* was defined as greater than or equal to 98% homology to the closest germline match analyzed via DNA sequencing. Greater than 20% ZAP-70 expression was considered positive (high-risk)[2]. Cytogenetics were evaluated by FISH for the most common CLL abnormalities (*del*(13q),

tri(12), *del*(11q), *del*(17p), rearrangements of chromosome 14) (probes from Vysis, Des Plaines, IL). Samples were scored positive for a chromosomal aberration based on consensus cytogenetic scoring[3]. Statistical analysis considered the hierarchical Doehner classification[4].

DNA quality control: We used standard Broad Institute protocols as previously described[5,6]. Tumor and normal DNA concentration were measured using PicoGreen® dsDNA Quantitation Reagent (Invitrogen, Carlsbad, CA). A minimum DNA concentration of 5 ng/µl was required for sequencing. All Illumina sequencing libraries were created with the native DNA. The identities of all tumor and normal DNA samples were confirmed by mass spectrometric fingerprint genotyping of 95 common SNPs by Fluidigm Genotyping (Fluidigm, San Francisco, CA).

Whole exome sequencing (WES): Informed consent for genomic analyses of patients' samples was obtained prior to the initiation of sequencing studies. The study was approved by the Ulm University Ethics Committee, IRC/EC number 138/03. DNA was extracted and purified from CD19+ and CD19- fractions of PBMC for tumor and matched germline DNA, respectively, using the Qiagen all-prep kit (Quiagen, Hilden, Germany) according to its unmodified protocol. Libraries for whole exome sequencing were constructed and sequenced on either an Illumina HiSeq 2000 or Illumina HiSeq 2500 using 76 bp paired-end reads. Details of whole-exome library construction have been described elsewhere[7]. Standard quality control metrics, including error rates, percentage-passing filter reads, and total Gb produced, were used to characterize process performance before downstream analysis. Note that due to a change in the Agilent capture bait set, the *NOTCH1* hotspot was not covered in samples GCLL-199 through GCLL-313, which included 10 samples with *NOTCH1* c.7544_7545delCT deletions by Sanger

sequencing[8]. CLL samples found to harbor *NOTCH1* mutations by Sanger sequencing were subsequently submitted to targeted sequencing with Illumina TruSeq Custom Amplicon library, and sequenced on an Illumina MiSeq with a mean coverage depth of 1332X. In the relapse samples, samples with *NOTCH1* mutations by Sanger sequencing (GCLL-0146-T-02, GCLL-0192-T-02, GCLL-0049-T-02, GCLL-0208-T-02) were also submitted to targeted sequencing.

In addition to samples from patients included in the CLL8 clinical trial, we have analyzed 157 WES samples from the cohort we have previously published[9]. The sequencing reads were realigned to hg19 and all downstream analysis was done using the same methods as with the sequencing data for the CLL8 cohort. Finally, previously published WES data for 103 matched CLL and germline DNA samples were downloaded with permission from the European Genome-Phenome Archive[10]. The raw sequencing reads were processed in identical fashion to the in-house produced WES libraries. New WES data is deposited in dbGaP (phs000922.v1.p1).

Identification of somatic mutations: Output from Illumina software was processed by the "Picard" data processing pipeline to yield BAM files containing aligned reads (bwa version 0.5.9, to the NCBI Human Reference Genome Build hg19) with well-calibrated quality scores[5,11]. From the sequencing data, somatic alterations were identified using a set of tools within the "Firehose" pipeline, developed at the Broad Institute (www.broadinstitute.org/cancer/cga). The details of our sequencing data processing have been described elsewhere[5,6]. Somatic single nucleotide variations (sSNVs) were detected using MuTect (Firehose version v131[12]); somatic small insertions and deletions were detected using an improved version (manuscript in preparation, Cibulskis et al.,) of Indelocator[5]. The primary improvement is implementation of local reassembly, which results in more accurate allele

fraction estimation. Following our standard procedure, we filter sSNVs and sINDELs by removing events seen in sequencing data of a large panel of normal samples. Overall, this filtering removed ~35% of all candidate somatic mutations, mostly ones with very low allelic fraction. In order to ensure that no candidate driver mutation were mistakenly removed by the filter, after completing the MutSig process, all filtered events in candidate CLL genes were also manually reviewed using the integrated genome viewer (IGV)[13]. In addition all mutations in candidate CLL genes were confirmed by manual inspection as well. The Oncotator tool was used to annotate mutations[14]. Conservation across 46 vertebrate species was performed and scored as previously described[9]. Sample contamination by DNA originating from a different individual was assessed using ContEst [15]. Median contamination value was 0.1% [inter-quartile range 0.1-0.3%]. Ig loci mutations were not included in this analysis.

In the 59 longitudinal samples, we utilized "forced calling" to quantify the number of alternate and reference alleles at sites with somatic mutation detected in a different sample from the same patient which was taken at a different time point, using the Samtools suite. Reads were considered if they were not marked as duplicate reads, had a base quality score at the site of interest ≥ 20 and alignment quality score ≥ 5, based on GATK BQSR base quality recalibration[11].

Estimation of and correction for tumor in normal content: For samples from the CLL8 trial, the majority of matched germline DNA samples were obtained from the CD19- fraction of PBMC. We found that some of these germline samples contained substantial proportions of tumor DNA, which could significantly decrease the ability to detect somatic mutations with high sensitivity using MuTect. We thus applied deTiN (manuscript in preparation, Taylor-Weiner et

al.), a method for estimating the level of tumor cells in the normal paired sample and recovering somatic mutations that would otherwise be filtered out due to evidence of the mutation in the normal. In brief, we estimate the level of tumor-in-normal (TiN) using two complementary approaches: (i) for tumors with a sufficient number of somatic mutations (total number of sSNVs and sINDELs >5), we use a linear fit of the respective tumor and normal allele fractions of candidate somatic mutations; and (ii) for tumors with sufficiently large sCNVs, we fit the allele frequency shift of germline heterozygous SNPs to a mixture of tumor and normal which provides an independent estimate for TiN. Next we use the TiN estimate to recover sSNVs and sINDELs that are at least a 1000 times more likely to be a somatic mutation than a germline event.

Significance analysis for recurrently mutated genes: We used MutSig2CV[16] to detect candidate cancer genes using three signals of positive selection: (i) increased mutation burden as compared to a background model; (ii) clustering of mutations along the gene; and (iii) enrichment of mutations at likely functional sites.

Genome-wide copy number analysis: Genome-wide copy number profiles of the CLL samples and their patient-matched germline DNA were estimated directly from the WES data, based on the ratio of CLL sample read-depth to the average read-depth observed in normal samples for that region. We observed a high level of agreement between sCNV detection by exome and standard FISH cytogenetics, with the exception of smaller deletions in the region of chromosome 13q14, where 14.5% of cases were missed by WES (***Supplementary Table 5***). Allelic copy-number analysis was then performed by examination of alternate and reference read counts at heterozygous SNP positions (as determined by analysis of the matched normal sample). These

counts were used to infer the contribution of the two homologous chromosomes to the observed copy-ratio in each segment. Further analysis of change-points in these allelic-ratios was performed using PSCBS[17], refining the segmentation. Finally, for each segment, we combined the copy-ratio and allelic data to derive allelic copy-ratios.

Significant recurrent chromosomal abnormalities were identified using the GISTIC2.0 algorithm[18] (v87). Regions with germline copy number variants were excluded from the analysis. Arm level and focal deletion and amplifications filtered by FDR Q<0.01 for significance. We identified 8 recurrent somatic arm level events in 157 of 538 patients (***Fig. 1***). GISTIC2.0[19] analysis yielded 4 significant arm level amplifications, including the previously described amplification of chromosome 8q (n=15) and trisomies of chromosomes 12 (n=72), 2p (n=47), and 19 (n=6)[20,21]. We also identified a significant focal amplification peak at 2p15 with 16 genes, which includes the CLL driver gene *XPO1*. We noted that *XPO1* mutations and amplifications involving the 2p arm were mutually exclusive in our cohort. Recurrent arm level deletions included *del(*17p) (n=34), *del(*8p) (n=19) , *del*(18p) (n=13) and *del*(20p) (n=7) as previously described[20,22]. Focal deletions included the canonical *del*(13q14.2) (n=255, containing *mir-15a* and *mir-16-1*), *del*(11q22.3) (n=118, containing *ATM*), a large deletion in 6q21 (n=18) (with a peak region spanning 72MB)[23], as well as expected deletions in the immunoglobulin (Ig) loci (***Supplementary Table 6*** contains a list of genes in the minimal deleted or amplified regions).

RNA sequencing and data analysis: RNA sequencing (RNAseq) was performed as previously described[24]. In addition, previously published RNAseq data for additional CLL RNA samples were downloaded with permission from the European Genome-Phenome Archive[25], and processed in an identical fashion to the in-house produced libraries. In total, matching WES and

RNAseq data were available for 156 samples including 103 samples collected at the DFCI and 53 samples collected by the ICGC. RNAseq BAMs were aligned to the hg19 genome using the TopHat suite. Each somatic base substitution detected by WES was compared to reads at the same location in RNAseq. Based on the number of alternate and reference reads, a power calculation was obtained with beta-binomial distribution (power threshold used was greater than 90%). A mutation call was deemed negative if no alternate allele reads were observed in RNA-Seq at the site, as long as RNAseq was powered to detect an event at the specified location.

Differential gene expression analysis was performed for the novel two driver gene mutations which affect at least 5 samples with matched RNAseq and WES data (*RPS15* [3 DFCI samples and 2 ICGC samples] and *MGA* [4 DFCI samples and 2 ICGC samples). Gene expression in transcripts per million (TPM) was quantified using the RSEM algorithm[26] (v1.2.19). Significant batch effects were seen between the DFCI and ICGC samples and were addressed as described below. After filtering non-expressed genes, gene expression was compared between samples with mutations in one of the studied driver genes and samples that are wild type for this gene. To address the significant batch affected, a generalized linear model (glm) was applied that includes both the mutation status and the batch information; and the *P* values for the mutation status coefficients were subjected to FDR correction (Q<0.1, listed in ***Supplementary Table 4***). Gene set enrichment analysis was performed using the GSEA software[27] (v2.2.0), with the pre-ranked list option. Genes were ranked based on the log transformed *P* values of the glm mutation coefficients.

Co-occurrence analysis: First, we considered two potential important confounders: prior therapy and *IGHV* mutation status, which may affect the proportion of patients affected by

specific drivers (***Extended Data Fig. 6A-B***) resulting in spurious instances of significant low or high co-occurrence. Indeed, despite similar average numbers of coding mutations per sample (24.3 +/- 11.2 versus 23.0 +/- 13.4 in *IGHV* mutated vs. unmutated samples, P=0.246), the median number of driver mutations per sample was higher in the unmutated *IGHV* subtype (3 [IQR 2-4] vs. 1 driver per sample [IQR 1-2], rank sum P<0.00001, note that these do not include *IGHV* as a driver, as *IGHV* mutations may represent a physiologic process in B cells). Similarly, compared to treatment-naïve cases, prior exposure to treatment was associated with an increased average number of coding sSNVs and sINDELs (34.9 +/- 22.7 vs. 23.0 +/- 11.4, $P = 1.16 \times 10^{-7}$), and a higher median number of drivers (4 vs. 2, $P = 3.55 \times 10^{-9}$). Of note, to examine the explanation that higher number of mutations simply reflects a longer time from diagnosis to sampling of patients with prior therapy, linear regression model analysis to evaluate the impact of the time between diagnosis and sample acquisition on the association between prior therapy and the number of mutations and drivers was performed. We found that in a model that included both time from diagnosis to sampling and prior therapy, only prior therapy retained significance in terms of a positive association with the number of mutations (P = 0.27147 and P=0.00634, respectively) and in terms of a positive association with the number of drivers (P = 0.201 and P= 1.66e-10, respectively). Therefore, to address the effect of *IGHV* and prior therapy status, significant low or high co-occurrence patterns were retained if the combined P value[28] for the tests in the two subsets after multi-hypotheses correction was significant at Q<0.1.

Estimation of mutation cancer cell fraction using ABSOLUTE: We used the ABSOLUTE algorithm (v1.1) to calculate the purity, ploidy, and absolute DNA copy-numbers of each sample[29], as previously described[9]. sCNVs were classified as clonal if the modal CCF estimate

exceeded 0.85. Modifications were made to the algorithm for the purpose of assessing the clonality of sSNVs and sINDELs as follows. For each mutation, the CCF probability density is estimated based on the mutation reference and alternate allele counts (t_ref_count and t_alt_count, respectively), the tumor purity, and local copy number for each homologous allele. The first step is to calculate the allele fraction probability distribution for the tumor alone, excluding the contribution from the normal fraction of cells. The proportion of tumor DNA at a site is:

$$Tumor\_DNA\_fraction = (purity \bullet CN_T) / (purity \bullet CN_T + CN_N (1-purity))$$

Where CN_T is the local copy number in the tumor cell and CN_N is the local copy number in the normal cells (2 in the autosome, and 1 or 2 on the X chromosome depending on gender). The allele fraction probability density in the tumor is estimated from the binomial probability density "*binopdf*" over the range of reference allele counts in the tumor "*t_ref*" between 0 and t_ref_count:

$$w(t\_ref) = binopdf(t\_ref + t\_alt\_count, t\_ref\_count + t\_alt\_count, Tumor\_DNA\_fraction)$$

The probability distribution for the allele fraction in the tumor is then:

$$p(AF_T) = \Sigma_{t\_ref} (w(t\_ref) \bullet betapdf(AF_T, t\_alt\_count+1, t\_ref+1))/Z$$

Where AF_T is the full range of allele fractions from 0 to 1 and Z is set such that $p(AF_T)$ is normalized to 1.

At this point, $p(AF_T)$ no longer contains the normal cell component, which simplifies the remaining steps to estimate the mutation cancer cell fraction (*CCF*). The CCF estimate is integrated over all possible mutation multiplicities "*m*", the number of mutations per tumor cell. The multiplicity m can range from 1 to q_hat_1 or q_hat_2 (the local somatic absolute copy

numbers of each homologous allele in the tumor). We assume that the mutation occurred before or after any local copy number change and each possible multiplicity is given equal weight w_m such that the w's are normalized to 1. AF_T is transformed to CCF coordinates by $CCF = AF_T \bullet CN_T/m$ and the probability density for CCF's ranging from 0 to 1:

$$p(CCF) = \Sigma_m (w_m \bullet p(AF_T)) \text{ for } AF_T \bullet CN_T/m < 1$$

and for $AF_T \bullet CN_T/m > 1$, the probability density was accumulated at $CCF=1$:

$$p(CCF=1) = \Sigma_m (w_m \bullet p(AF_T \bullet CN_T/m)) \text{ for } AF_T \bullet CN_T/m >= 1$$

Clonal mutations were defined as sSNVs or sINDELs with p(CCF>0.85)>0.5 (ie. median CCF. greater than 0.85).

Clustering analysis of sSNVs and sINDELs in 59 CLL sample pairs: We performed WES on matched samples collected at the time of first progression following therapy from 59 of 278 CLL8 subjects (***Supplementary Table 9 & 10*****).** The median time to progression was 35.1 months (range 5.9-75.5), with relapse samples collected at a median of 7.6 months following documented progression, all before receipt of subsequent therapy. The two time point CCF clustering procedure was performed as previously described[9]. Clonal evolution between pre-treatment and relapse samples was defined based on the presence of mutations with a $P(\Delta \text{ CCF} \geq 0.1)>0.5$. Branched evolution was classified when a dominant clone in the pre-treatment sample was replaced by sibling dominant clone. This pattern was indicated by a CCF decrease of the mutations in the pre-treatment dominant clone co-occurring with a CCF increase of mutations in the relapse dominant clone. In contrast, linear evolution (replacement of a parent clone by its progeny) was indicated when the increase in CCF of the relapse dominant clone was not

accompanied by a decrease in CCF of the pre-treatment dominant clone.

For the CLL driver analysis (***Fig. 5C***), a significant change in CCF over time (red or blue) was determined if the 95% CIs of the CCF in the pre-treatment and relapse sample did not overlap. For each driver, a binomial test was performed to assess whether the proportion of instances within each category (increases, decreases, stable) significantly exceeded 0.5 (with a BH FDR correction, $Q<0.1$).

Deep sequencing of somatic single nucleotide variants

Targeted deep sequencing was performed using microfluidic PCR (Access Array System, Fluidigm). Six unmatched saliva samples were included in this analysis to assist with the quantification of background sequencing error noise. Target-specific primers were designed to flank sites of interest and produce amplicons of 200 bp ± 20 bp. Per well, molecularly barcoded, Illumina-compatible specific oligonucleotides containing sequences complementary to the primer tails were added to the Fluidigm Access Array chip together with genomic DNA samples (20–50 ng of input) such that all amplicons for a given DNA sample shared the same index, and PCR was performed according to the manufacturer's instructions. From each individual collection well from the Fluidigm chip, indexed libraries were recovered for each sample, quantified using picogreen, and then normalized for uniformity across libraries. Resulting normalized libraries were loaded on the MiSeq instrument and sequenced using paired-end 150 bp sequencing reads[30]. We confirmed the presence of a mutation in a sample if the fraction of alternate reads exceeded that in the normal control samples (beta binomial test, FDR $Q<0.1$).

Order of mutations analysis: To investigate the question of temporal ordering of driver appearance in CLL, the driver clonality patterns in the 501 of 538 samples which were treatment-naïve were studied using the following approach: Whenever a driver event d1 is clonal and another driver event d2 is subclonal in the same sample, this pattern indicates that d1 was acquired before d2 (denoted by d1 → d2), and allow us to draw an edge between these two drivers (out-going for d1, and in-going for d2). Assuming the temporal ordering of a driver pair is random, we can apply hypothesis testing to find significant temporal orderings among drivers in this data set[31]. We classified driver events as early, late or intermediary/not powered by counting the in-degrees and out-degrees of each driver event across the 501 samples and applying a two-tailed binomial test to quantify whether a driver event has a significantly greater number of out-degrees (early), a significantly greater number of in-degrees (late) or no significant preference (intermediary/not powered). To account for multiple hypothesis testing, we additionally calculate the corresponding q-values as a measure of significance in terms of the false discovery rate[32].

To infer a temporal order between any two pairs of drivers, all known CLL driver pairs were considered for which at least 5 clonal-subclonal orderings were observed, as this was the minimal number of observations powered for statistical significance ($P<0.1$) to detect a completely unidirectional relationship (i.e., d_1 is clonal and d_2 is subclonal in all five pairing). Two-tailed binomial tests were used to quantify the confidence in the temporal ordering of each pair of driver mutations. To account for multiple hypothesis testing, the corresponding q-values were further calculated as a measure of significance in terms of the false discovery rate[32].

Statistical methods Statistical analysis was performed with MATLAB (MathWorks, Natick, MA), R version 2.11.1 and SPSS version 21 (IBM, NYC, NY). Categorical variables were compared using the Pearson Chi-square test or Fisher Exact test as appropriate, and continuous variables were compared using non-parametric rank-sum tests. Statistical analyses of data from the CLL8 clinical trial were performed on an intention-to-treat basis meaning that all eligible subjects with available samples were analyzed as randomized. Time to event analyses were done for progression free survival (PFS), which was defined as the time from randomization to disease progression or death, and for overall survival (OS), which was understood as the time between randomization and death. Time to event data were estimated by the Kaplan–Meier method, and differences between groups were assessed using two-sided non-stratified log-rank tests. Additionally, hazard ratios (HR) and 95% confidence intervals (CI) were calculated using unadjusted and adjusted Cox regression modeling. Independent factors for PFS and OS were identified by multivariable analysis using Cox proportional hazards regression models.

SUPPLEMENTARY BIBLIOGRAPHY

1 Hallek, M. *et al.* Addition of rituximab to fludarabine and cyclophosphamide in patients with chronic lymphocytic leukaemia: a randomised, open-label, phase 3 trial. *Lancet* **376**, 1164-1174, doi:10.1016/S0140-6736(10)61381-5 (2010).

2 Rassenti, L. *et al.* Relative value of ZAP-70, CD38, and immunoglobulin mutation status in predicting aggressive disease in chronic lymphocytic leukemia. *Blood* **112**, 1923-1930, doi:blood-2007-05-092882 [pii]10.1182/blood-2007-05-092882 (2008).

3 Smoley, S. A. *et al.* Standardization of fluorescence in situ hybridization studies on chronic lymphocytic leukemia (CLL) blood and marrow cells by the CLL Research Consortium. *Cancer Genet Cytogenet* **203**, 141-148, doi:10.1016/j.cancergencyto.2010.08.009 (2010).

4 Döhner, H. *et al.* Genomic aberrations and survival in chronic lymphocytic leukemia. *N Engl J Med* **343**, 1910-1916, doi:MJBA-432602 [pii]10.1056/NEJM200012283432602 (2000).

5 Chapman, M. A. *et al.* Initial genome sequencing and analysis of multiple myeloma. *Nature* **471**, 467-472, doi:10.1038/nature09837 (2011).

6 Berger, M. F. *et al.* The genomic complexity of primary human prostate cancer. *Nature* **470**, 214-220, doi:10.1038/nature09744 (2011).

7 Fisher, S. *et al.* A scalable, fully automated process for construction of sequence-ready human exome targeted capture libraries. *Genome Biol* **12**, R1, doi:10.1186/gb-2011-12-1-r1 (2011).

8 Stilgenbauer, S. *et al.* Gene mutations and treatment outcome in chronic lymphocytic leukemia: results from the CLL8 trial. *Blood* **123**, 3247-3254, doi:10.1182/blood-2014-01-546150 (2014).

9 Landau, D. A. *et al.* Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell* **152**, 714-726, doi:10.1016/j.cell.2013.01.019 (2013).

10 Quesada, V. *et al.* Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nat Genet* **44**, 47-52, doi:10.1038/ng.1032 (2012).

11 DePristo, M. A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* **43**, 491-498, doi:10.1038/ng.806 (2011).

12 Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nature Biotechnol* **31**, 213-219, doi:10.1038/nbt.2514 (2013).

13 Robinson, J. T. *et al.* Integrative genomics viewer. *Nature Biotechnol* **29**, 24-26, doi:10.1038/nbt.1754 (2011).

14 Ramos, A. H. *et al.* Oncotator: Cancer Variant Annotation Tool. *Hum Mutat*, doi:10.1002/humu.22771 (2015).

15 Cibulskis, K. *et al.* ContEst: estimating cross-contamination of human samples in next-generation sequencing data. *Bioinformatics* **27**, 2601-2602, doi:10.1093/bioinformatics/btr446 (2011).

16 Lawrence, M. S. *et al.* Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* **505**, 495-501, doi:10.1038/nature12912 (2014).

17 Olshen, A. B. *et al.* Parent-specific copy number in paired tumor-normal studies using circular binary segmentation. *Bioinformatics* **27**, 2038-2046, doi:10.1093/bioinformatics/btr329 (2011).

18 Mermel, C. H. *et al.* GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol* **12**, R41, doi:10.1186/gb-2011-12-4-r41 (2011).

19 Mermel, C. H. *et al.* GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol* **12**, R41, doi:10.1186/gb-2011-12-4-r41 (2011).

20 Edelmann, J. *et al.* High-resolution genomic profiling of chronic lymphocytic leukemia reveals new recurrent genomic alterations. *Blood* **120**, 4783-4794, doi:10.1182/blood-2012-04-423517 (2012).

21 Sellmann, L. *et al.* Trisomy 19 is associated with trisomy 12 and mutated IGHV genes in B-chronic lymphocytic leukaemia. *Br J Haematol* **138**, 217-220, doi:BJH6636 [pii]10.1111/j.1365-2141.2007.06636.x (2007).

22 Rudenko, H. *et al.* Characterising the TP53-deleted subgroup of chronic lymphocytic leukemia: an analysis of additional cytogenetic abnormalities detected by interphase fluorescence in situ hybridisation and array-based comparative genomic hybridisation. *Leuk Lymphoma* **49**, 1879-1886, doi:904754482 [pii] 10.1080/10428190802345902 (2008).

23 Cuneo, A. *et al.* Chronic lymphocytic leukemia with 6q- shows distinct hematological features and intermediate prognosis. *Leukemia* **18**, 476-483, doi:10.1038/sj.leu.2403242 (2004).

24 Landau, D. A. *et al.* Locally disordered methylation forms the basis of intratumor methylome variation in chronic lymphocytic leukemia. *Cancer Cell* **26**, 813-825, doi:10.1016/j.ccell.2014.10.012 (2014).

25 Ferreira, P. G. *et al.* Transcriptome characterization by RNA sequencing identifies a major molecular and clinical subdivision in chronic lymphocytic leukemia. *Genome Res* **24**, 212-226, doi:10.1101/gr.152132.112 (2014).

26 Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC bioinformatics* **12**, 323, doi:10.1186/1471-2105-12-323 (2011).

27 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *PNAS* **102**, 15545-15550, doi:10.1073/pnas.0506580102 (2005).

28 Whitlock, M. C. Combining probability from independent tests: the weighted Z-method is superior to Fisher's approach. *J Evol Biol* **18**, 1368-1373, doi:10.1111/j.1420-9101.2005.00917.x (2005).

29 Carter, S. L. *et al.* Absolute quantification of somatic DNA alterations in human cancer. *Nature Biotechnol* **30**, 413-421, doi:10.1038/nbt.2203 (2012).

30 Lohr, J. G. *et al.* Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *PNAS* **109**, 3879-3884, doi:10.1073/pnas.1121343109 [pii] (2012).

31 Wang, J. *et al.* Tumor evolutionary directed graphs and the history of chronic lymphocytic leukemia. *eLife* **3**, doi:10.7554/eLife.02869 (2014).

32 Storey, J. D. & Tibshirani, R. Statistical significance for genomewide studies. *PNAS* **100**, 9440-9445, doi:10.1073/pnas.1530509100 (2003).